AD_____

Award Number: DAMD17-00-1-0244

TITLE: Improving Breast Cancer Research Through Automated Matching of Patients to Clinical Trials

PRINCIPAL INVESTIGATOR: Lawrence O. Hall, Ph.D.
Dmitry B. Goldgof, Ph.D.
Jeffrey Krischer, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, Florida 33620-7900

REPORT DATE: August 2001

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20011128 199

			Form Approved
REPORT DOCUMENTATION PAGE			OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE August 2001	3. REPORT TYPE AND DATES COVERED Annual (3 Jul 00 - 2 Jul 01)	
4. TITLE AND SUBTITLE Improving Breast Cancer Research Through Automated Matching of Patients to Clinical Trials			5. FUNDING NUMBERS DAMD17-00-1-0244
6. AUTHOR(S) Lawrence O. Hall, Ph.D. Dmitry B. Goldgof, Ph.D. Jeffrey Krischer, Ph.D.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of South Florida Tampa, Florida 33620-7900 E-Mail: hall@csee.usf.edu			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited	12b. DISTRIBUTION CODE

13. Abstract (Maximum 200 Words) (abstract should contain no proprietary or confidential information)

A Web based prototype intelligent agent/expert system for matching breast cancer patients to clinical trials has been built. It allows for cost preferences to be entered. Therefore, the system user can choose to rule patients out of trials as quickly as possible without regard to the cost of tests necessary to do this. They can choose have questions appear so that the patient is ruled out of the trial with the minimal set of costs (tests) or can choose some combination of approaches. The system has been tested with seven protocols and designed for maximal responsiveness and scalability as new protocols are added. The files of 146 former patients have been used to test the accuracy of the system. Patients for each of the seven protocols were correctly found eligible for one or more trials.

14. SUBJECT TERMS cancer control, clinical trials, expert system			15. NUMBER OF PAGES 11
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Contents

3 Introduction

Increasing the enrollment of patients in clinical trials is important to making progress towards finding more effective treatments for breast cancer. Accrual is complicated by a large number potential studies and the cost and complexity of determining whether a patient meets the necessary eligibility criteria. Under this proposal, we are developing a Web based expert system which can determine the patients eligibility for clinical trials. The expert system is designed to take into account the cost of tests which are required to meet inclusion criteria and acquire information in the most cost-effective way possible.

4 Body

In this first-year, we have built an expandable prototype system. There are currently seven breast cancer protocols available in the system. They are shown in Figure 1 using their Moffitt Cancer Center number and a brief summary description. In a retrospective study, the records of 125 patients who were assigned to one of the trials were used to provide data for prototypical expert system to make clinical trial assignments. An additional 21 records of patients who were assigned to trials that are not implemented within our system were tested for eligibility for the 7 implemented protocols/trials.

The results of determining the eligibility for the 146 patients are reported in four categories. They are the true positives (patients that the system found were eligible for a clinical trial and they were assigned to that trial), the false positives (patients that the system found were eligible for clinical trial and were not assigned to that clinical trial), false negatives (patients that the system found were ineligible for a trial and they were assigned to that trial), and true negatives (patients that the system found were ineligible for a trial and they were not assigned to that trial). The results from our prototypical system are shown in Table 1. Each false positive and false negative was separately compared against all inclusion and exclusion criteria to make sure the patient was correctly determined to be eligible or ineligible for particular protocol. All other results matchup with actual patient assignments to clinical trials.

The false positive category is interesting because these patients were eligible for multiple trials. Not all of the protocol/trials implemented by our system were available at the same time. Since the data is retrospective, it is impossible to tell whether a physician considered a patient for more than one trial and then chose the best trial or simply chose a trial they were familiar with for which the patient was eligible. However, it is interesting that the system finds 43 of 146 cases in which patients were eligible for multiple trials. This suggests that the expert system will help prevent a viable trial from being overlooked.

The false negatives all occurred due to some lack of data in the patient chart. For example, the number of positive lymph nodes may not have been found by us. There was no case in which the patient was assigned to a trial but clearly violated an inclusion criteria or met an exclusion criteria.

The developed prototype has a cost function built into it. It can be required to order questions so that the tests associated with them are of increasing cost amounts. We also have the pain cost associated with each test which is in turn associated with one or more questions. Questions can be ordered to cause the least amount of pain. Combinations of monetary and pain costs

and number of questions to determine inclusion/exclusion are also possible for question ordering. In the absence of a requirement for cost minimization, a page of approximately 10 questions is presented with the questions ordered on the page by how likely the question is to rule in or rule out a patient from one or more protocols. At all times, the user of the system can answer one, several, or all of the questions on a page before submitting their answers. Figure 2 shows an initial page of common questions that is displayed as the initial question page.

At the present time, we are in the process of testing the cost functionality. There are still some questions for which we do not have the cost of the tests associated with them. This problem is to be resolved this week. In initial experiments utilizing monetary cost, we were usually able to exclude a patient with less cost in the case that a user answered questions in order. It is possible, for example, that a series of three $40 tests might be required as a group to answer relevant questions and the questions associated with them would the individually ordered before a question associated with a single $100 test that might have similar impact in determining eligibility/ineligibility. This could cause it to be more expensive to use cost functionality as currently implemented in this kind of special case. As part of our future research, an investigation into grouping sets of related tests so that their cost is considered as a single cost will be undertaken. We have not yet experimented with pain cost.

The opening page of version 1.2 of the system is shown in Figure 3. In Figure 4, a status page after number of questions has been answered is shown. It is possible to ask why a patient was excluded from protocol and, if necessary, change answers which could change the result. We are in the process of adding the why button for patients who are assigned to a trial. The prototype system is available online at http://morden.csee.usf.edu/moffit and a password is available from the principal investigator.

At the present time, new protocols must be entered into the system by a computer programmer. We have developed the companion system which enables a nurse, technician or physician to enter a protocol. There is a not completely working prototype done. It matches questions with tests to attempt to have just one question associated with the need for results from a specific test so that the user does not get asked "the same type of questions" multiple times. This will be a focus of our work in year 2.

In year 2, we will completely debug and utilize cost functionality, further refine the prototype for automating the new protocol addition process, and a new set of current and active breast cancer protocols/trials will be added. We will get feedback on the user interface from physicians and refine it to make the system as pleasant to use as possible. After developing a good user interface, we want to start working with prospective patients who may be assigned to clinical trials/protocols. This research will be done in parallel to current clinical practice without influencing it.

Key Research Accomplishments:

- A prototype (version 1.2) of a Web based expert system for assigning patients to clinical trials has been developed and tested. It correctly assigns 146 patients to clinical trials and a retrospective study.

- Utilizing retrospective patient data, it has been found that patients are eligible for multiple protocols/trials.

The Clinical Trial Assignment Expert System: Patient Entry (Version 1.2)

Questions about this page? See instructions below.

Patient Name	
Patient ID Number	NEW PATIENT Click to enter a new patient PATIENT SEARCH Click if this patient was previously entered
CLEAR	

Instructions:
If the patient is being entered for the first time:

    1. Completely type in the name and assign an identification number.
    2. Click on "NEW PATIENT."

If the patient's information has been previously entered:

    1. Type either the patient's name or the ID number in the appropriate field.
    2. Click on "PATIENT SEARCH."

Figure 1: A list of implemented protocols.

The Clinical Trial Assignment Expert System: Initial Questions for BreastCancer (Version 1.2)

Patient Name: trial
Patient ID: 722001
Questions about this page? See instructions below.

1. What is the stage of the breast cancer? ($ 0.00)	2. What is the patients age in years? ($ 0.00)
3. What is the patients gender? ($ 0.00) Female Male	4. How many nodes are positive? (Enter 0 if there are none) ($ 0.00)
5. What is the greatest diameter of the tumor in cm? ($ 0.00)	6. What is the estrogen receptor status? ($ 80.00) Positive Negative Unknown
7. What is the progesterone receptor status? ($ 80.00) Positive Negative Unknown	8. Is the patient pregnant or nursing? ($ 0.00) Yes No Unknown
9. Is the patient considered a candidate for adjuvant or first-line hormonal therapy? ($ 0.00) Yes No Unknown	10. Has the patient had surgery for breast cancer? ($ 0.00) Yes No Unknown
11. If the patient has had surgery for breast cancer, what was the most extensive type? ($ 0.00)	12. Did the surgery include an axillary dissection? ($ 0.00) Yes No Unknown
13. Has the patient been administered any therapy for cancer? ($ 0.00) Yes No Unknown	14. If the patient was administered therapy for cancer, what types were used? ($ 0.00) Hormonal Radiation Chemotherapy Immunotherapy
PROCESS Click to submit your answers to the system	

Figure 2: The initial page of questions.

Welcome to the Expert System for Clinical Trial Assignment: (Version 1.2)

This expert system has been designed to help identify clinical trials for which your patient may be eligible. You will be asked to answer questions related to protocol eligibility criteria and the system will determine the eligibility of your patient for currently active clinical trials.

What is an expert system?

How the expert system works

Step-by-step instructions

ENTER Click to enter the clinical trial assignment system.

Figure 3: Opening page of the Expert System.

The Clinical Trial Assignment Expert System: Protocol Questions (Version 1.2)

Patient Name: trial
Patient ID number: 722001

Protocol (Find out about each Protocol)	Status	Maximum Questions Remaining	Percentage of Questions Answered
10822 (Subject)	NOT ELIGIBLE Why?	0	100%
11378 (Subject)	NOT ELIGIBLE Why?	0	100%
10840 (Subject)	NOT ELIGIBLE Why?	0	100%
11072 (Subject)	MORE INFORMATION NEEDED	33	5%
12100 (Subject)	MORE INFORMATION NEEDED	8	0%
12101 (Subject)	NOT ELIGIBLE Why?	0	100%
11992 (Subject)	NOT ELIGIBLE Why?	0	100%

Figure 4: Status listing during use of the Clinical Trial Assignment Expert System.

Table 1: Results from testing with 146 retrospective patients. TP - True Positive, FP - False Positive, FN - False Negative, TN - True Negative.

Protocols	12101	11072	10822	12100	11992	10840	11378
TP	20	48	10	8	5	0	4
FP	11	21	1	3	6	0	1
FN	0	19	0	13	0	3	3
TN	115	58	135	122	135	143	138

- Utilizing monetary cost in the expert system leads to a set of questions which usually result in less tests being needed to rule out a patient for a clinical trial, which results in less cost needed to make a determination.

- Significant progress has been made in the development of automated protocol acquisition tool. It is a specific type of knowledge acquisition tool.

Reportable Outcomes: Paper under preparation which describes the operation, underlying algorithms for question display, and capabilities of the system. Web prototype available at http://morden.csee.usf.edu/moffit with password available from the principal investigator.

5 Conclusions

We have developed a scalable prototype which currently can determine eligibility for seven breast cancer clinical trials. The system has been tested using retrospective data from 146 patients who are assigned to some clinical trial. Its accuracy has been verified. The system correctly finds cases (43) in which a patient is eligible for multiple clinical trials. This will enable a physician to make the best choice from available trials. The system is able to utilize monetary cost in requesting tests to rule in/rule out a patient from the set of available clinical trials. The default ordering of questions allows the system user to rapidly determine the eligibility or ineligibility of a patient for any subset of the available clinical trials entered into the system.

The system is Web based and password protected. It provides rapid response when a person enters answers to 1 or more questions on a page of system selected questions. It can be used from any computer on the World Wide Web. Hence, community physicians will be able to determine the potential eligibility (they may not wish to run all tests) of the patient for clinical trials at cancer centers in their region.

Most of the work on a prototype to enable physicians, nurses or technicians to enter new protocols has been completed. When this is complete and tested, it will allow for simple and fast addition of new protocols to the system. This knowledge acquisition tool is being designed to minimize/eliminate the cases where similar questions acquiring essentially the same information would have to be asked.

5.1 So What

The prototype system shows the potential for allowing community physicians, as well as cancer center physicians, to quickly and cost effectively determine for which clinical trials a patient may

be eligible. It holds the promise of enabling greater patient accrual for trials by increasing the awareness of each trial for treating physicians throughout a region.